Remarks

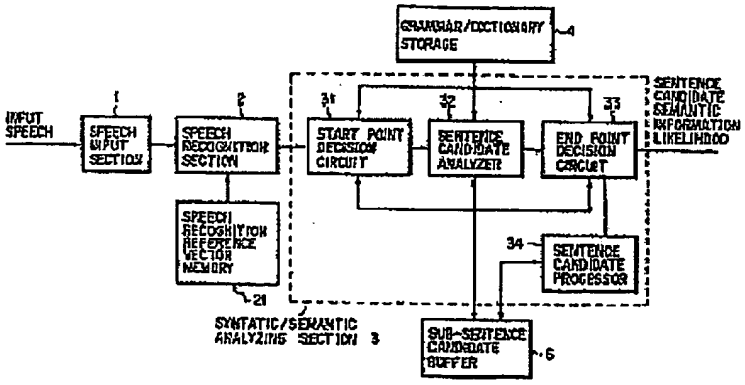
Claims 1-30 are pending in the present application, wherein original claims 1, 13, 23, 24, 26, and 29 are independent claims. Claim 1 has been amended for clarification. All claims have been variously rejected under 35 USC §102. Applicant traverses these rejections and requests withdrawal thereof.

***Rejections Under 35 USC. §102***

The Office Action rejects claims 1-30 under 35 USC §102(b) as being anticipated by U.S. Patent No. 5,457,768 to Tsuboi et al. ("'768 patent"). According to its Abstract, the '768 patent teaches that:

> A speech recognition apparatus comprises a speech input unit for receiving an input speech signal, analyzing it, and outputting a speech feature parameter series, a speech recognition unit for extracting a speech feature vector from the parameter series, and matching it with a plurality of predetermined words to output a series of word candidates used as keywords, a syntactic analysis unit for analyzing the series of the word candidates as the keywords according to syntactic limitation, and generating a sentence candidate.
>
> ('768 patent, Abstract, emphasis added)

FIG. 2 of the '768 patent includes the above units, as follows:




Whereas, the present invention takes the following form:

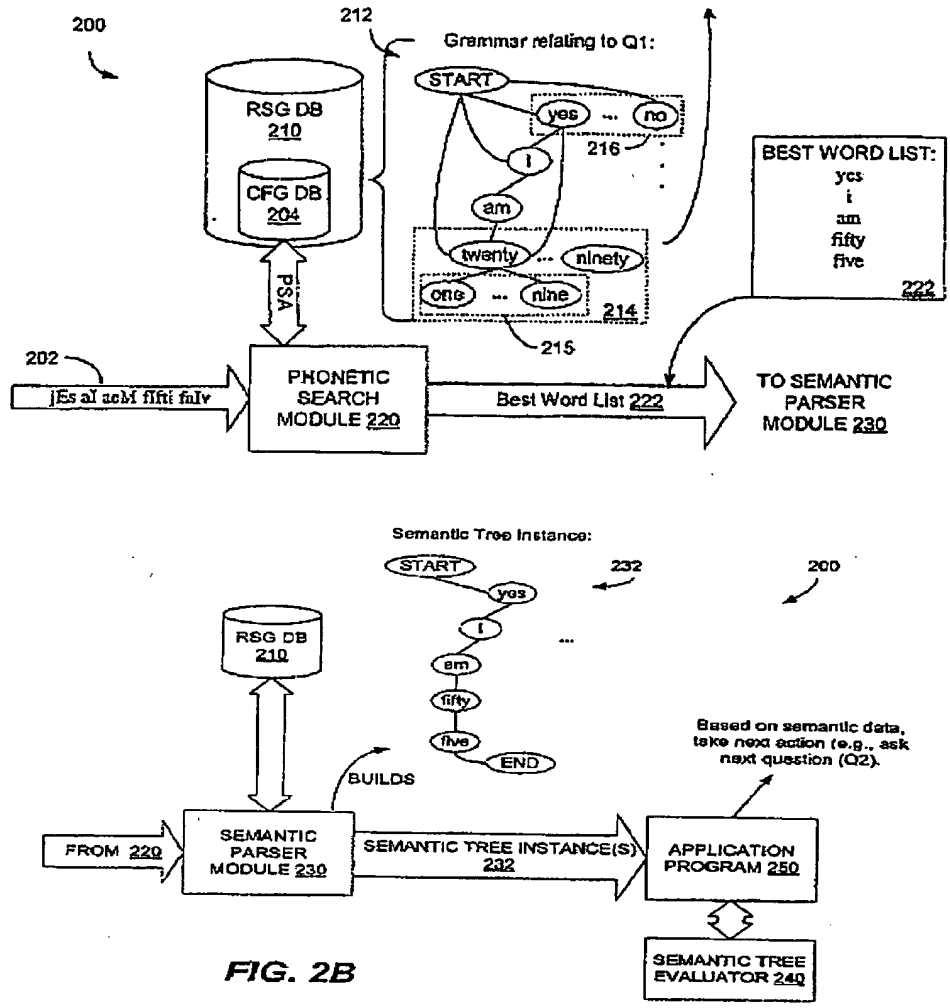



FIG. 2B

Claim 1

The office action asserts that each element of claim 1 is anticipated by the '768 patent. However, as shown by the below analysis, the '768 patent does not anticipate each and every element of claim 1.

The office action states that "Tsuboi teaches a phonetic data ("reference vectors", including words or phonemes) processing system (Fig. 17A, 22)", and then goes on to assert that the individual elements of claim 1 are anticipated by the '768 patent. Specifically, the Office Action appears to assert that element B of claim is anticipated by the '768 patent: "A context free grammar/dictionary

stored in memory (4) comprising syntactic and semantic information." The element B of claim 1 is as follows:

> B. a rich semantic grammar (RSG) stored in said memory and comprising one or more grammars comprising syntactic information and semantic information; and

This element of claim 1 is shown as RSG 210 in FIG. 2A and 2B of the present application, and reproduced in-part above. The text of the application supports the figures and the claim:

> The RSG serves as a common, root grammar for the syntactic and semantic processing, which may each use its own representation of the RSG for its respective processing tasks. As a result, compatibility is ensured, since the semantic stages of processing will never receive an illegal word string, for example, from the syntactic processing stage.
>
> (Application, p. 7, lines12-16)

As is shown in these figures, the RSG 210 is "common", in that it is used by both the *phonetic search module 220* (to generate a best word list) and by *semantic parser module 230* (to produce semantic tree instances). Yet, in the '768 patent two different databases are used: speech recognition reference vector memory 21 and grammar/dictionary storage 4. The *speech recognition unit* 2, which produces the "word lattice", uses the *reference vector memory 21*, but not the *grammar/dictionary storage 4*. The '768 patent's *syntactic/semantic analyzing section 3* (shown above), which receives the "word lattice" uses the *grammar/dictionary storage 4*, and not the *reference vector memory 21*. Suffice to say, the generation (and maintenance) and use of two different databases in the '768 patent does not anticipate a common, root grammar such as the RSG of claim 1.

There are further critical difference between the RSG and the reference vector memory 21 of the '768 patent content. The '768 patent describes:

> The speech recognition unit 2 is connected to a speech recognition reference vectors memory 21 ... storing speech feature patterns of keywords corresponding to predetermined words to be recognized ... and matches the speech feature parameter series with the reference vector in the speech recognition reference vector memory 21.
>
> ('768 patent, col. 4 lines 50-56)

> In order to understand the input speech, keywords needed for semantic expression are detected from the input speech, and syntactic and semantic analysis is performed to generate a semantic expression. The term "keywords" means "predetermined specific words needed for understanding a free speech to obtain a semantic expression".
>
> ('768 patent, col. 4 lines 25-30)

Thus, in the '768 patent the *reference vector memory 21* appears to consist of "keywords" – "predetermined specific words". And in the '768 patent, it is essential that the grammar includes predetermined "keywords". As such, they are heavily context dependent. In fact, as described, without detection of a keyword, absolutely no syntactic or semantic analysis will take place. In further contrast, there are no specific keywords "needed for understanding a free speech" in the RSG of the present invention – the RSG is not a database of "keywords", i.e., "predetermined specific words". Rather, in the present invention:

> The grammars of RSG DB are organized as a tree of classes, as nodes, that represent words and phrase structures. A grammar tree may include a plurality of trees (or sub-trees). Generally, the RSG DB is context-free, because the rules applied to each sub-tree operate independently of the other parts of the tree.
>
> (Application, p. 7, lines 19-22)

Furthermore, the *grammar/dictionary memory unit 4* of the '768 patent also uses "keywords", i.e., "predetermined specific words needed for understanding a free speech".

> FIG. 3 shows a word dictionary including keywords and examples of connectable word ranges (described below), and FIG. 4 shows examples of grammatical rules including rules for semantic analysis processing. These date items are stored in a grammar/dictionary memory unit 4.

The grammar/dictionary storage 4 includes the keywords, but as shown above the RSG of the present invention is not a composition of "keywords", as defined in the '768 patent. In fact, the grammar/dictionary of FIG. 3 of the '768 patent (i.e., WORD – PART OF SPEECH – CONNECTABLE RANGE) bears no resemblance with an RSG as used in the present invention, which take the form of a tree having nodes represented by classes having semantic and syntax

information. For all of the above reasons, Applicant believes that the common, root RSG of element B of claim 1 is not taught by the grammar/dictionary storage 4 of the '768 patent, nor is it taught by the reference vector memory 21.

The office action goes on to further assert that the '768 patent teaches "A 'phonetic' (word) estimating module (2) for producing a word list/ series, where each word of the word list/series has at least four data items ..." – presumably anticipating the "phonetic searcher" of claim 1, element C(1). The office action also states that "A word processing module (3), comprising a word searcher for generating sequences of word candidates from the grammar dictionary module, and generating sequences of word lattice;" Applicant points out that item (3) of the '768 patent (as shown in its FIG. 2, also above) is the syntactic/semantic analysis unit (3) – i.e., the unit that receives, but does not generate, a word lattice. The office action also asserts that the syntactic/semantic analysis unit (3) teaches element C(2) of claim 1. Applicant will address the above two statements in the office action, given the assumption they are meant to show the phonetic searcher of element C(1) of claim 1.

Element C(1) of claim 1 was amended slightly for clarification, and is restated below in clean form, but with emphasis added:

> C. a phonetic data processing module, executable by said processing device, said module comprising:
>
> (1) a *phonetic searcher*, configured to generate, using syntactic analysis and as a function of syntactic information derived from said RSG and a received phonetic stream comprised of phonetic estimates, a set of sequences comprising a set of best words from said RSG corresponding to said phonetic estimates;

In contrast, the *speech recognition unit 2* of the '768 patent generates word candidates in the form of a "word lattice",

> The recognition result of the speech recognition unit 2 is sent to a syntactic/semantic analysis unit 3 in the form of a series of temporally discrete words (i.e. word lattice). Each word of the word series has at least four data items, "start point", "end point", "likelihood" and "word name". The data item "likelihood", which indicates the possibility of words, may include "distance", "similarity" and "probability".

('768 patent, col. 5 lines 3-9)

As a first distinction, the *phonetic searcher* of this claim element uses a rich semantic grammar (RSG) to produce a *set of best words*. In contrast, the '768 patent teaches (i.e., see FIG. 2), a *speech recognition unit 2* that compares phonetic estimates from input speech with keyword vectors from *reference vector memory 21* to form a *word lattice*. Keywords have specific predetermined meanings relied on by the *speech recognition unit 2*. But the *phonetic searcher* of the present invention does not rely on "keywords" or use "keyword vectors", - it does not require any special predetermined words to generate a set of best words. Therefore, the *phonetic searcher* performs different functions to achieve a different result.

A further distinction exists between the *set of best words* and the *word lattice*. FIG. 8 of the '768 shows a word lattice produced by the speech recognition unit 2, which would then be passed to the syntactic/semantic unit 3, both shown in its FIG. 2 (depicted above). As is shown, a *word lattice*, by definition, has no connections or word paths.

> The recognition result of the speech recognition unit 2 is sent to a syntactic/semantic analysis unit 3 in the form of a series of temporally discrete words (i.e. word lattice).
>
> ('768 patent, col. 5 lines 3-5)

> The word spotting is thus performed, and a keyword series of word candidates of W1 to W6, as shown in FIG. 8, is obtained. The word candidates recognized in the speech recognition unit 2 are output to the syntactic/semantic analysis unit 3.
>
> ('768 patent, col. 8 lines 47-52)

Compare FIG. 8 of the '768 patent to FIG. 4D of the present application, which shows the connections and word paths for the set of best words (or best word list). See also the following text:

> In this manner, segments are combined and, when done, a resulting best word list 222 is formed that includes words and word paths, which include the segments that make up each sequence.
>
> (Application, p. 27, lines 3-5)

Thus, not only are the *search recognition section 2* of the '768 patent (and keyword vectors) and the *phonetic searcher* (and RSG) of the present invention functionally different, they produce

different results. Thus, for these additional reasons, it is believed that the '768 patent does not anticipate the present invention.

Finally, with respect to claim 1, the office action also appears to assert that the semantic parser module of claim 1, element C(2) is taught by the syntactic/semantic analyzing unit 3 of the '768 patent (see FIG. 2). This element of claim 1 is restated here, with emphasis added:

> C. a phonetic data processing module, executable by said processing device, said module comprising:
> ...
> (2) a *semantic parser module*, configured to generate a context free set of semantic data from said sequences and said RSG, wherein said set of semantic data includes all valid interpretations of the sequences.

In this element of claim 1, the *semantic parser module* generates context free semantic data using the common RSG – the same RSG used by the phonetic parser to produce the *set of best words*. But the *syntactic/semantic unit 3* of the '768 patent is different – it uses a *reference vector memory 21* to produce the word lattice and *grammar/dictionary 4* to produce semantic candidates. It is acknowledged that the grammar/dictionary also includes keywords, but the grammar/dictionary is structurally different from the reference vector memory 21. For example, in the '768 patent the word dictionary of FIG. 3 is much different from the keyword lattice shown in FIG. 8. Thus, the '768 patent teaches using 2 different databases of information to produce semantic data and the present invention teaches using one common RSG. For this reason, Applicant believes that the semantic parser of claim 1 is not anticipated by the '768 patent.

Additionally, as previously mentioned, the semantic parser module of claim 1 does not rely on "keywords", whereas the syntactic/semantic analyzing unit 3 of the '768 does require special predetermined keywords.

> The syntactic analysis method used in the syntactic/semantic analysis unit 3 is based on, e.g. the Tomita method (M. Tomita, "An efficient Word Lattice Parsing Algorithm for Continuous Speech Recognition", Proc. ICASSP '86, pp. 159-1572, Nov., 1986). In this method, however, the Tomita method is modified such that while the temporal connection relationship and syntactic connection relationship of temporal discrete keywords are analyzed, the semantic

> expression of the analyzed keyword series can be generated. Syntactic grammar/dictionary data are converted to an LR parse table in advance, and the syntax of the keyword series can be quickly analyzed by referring to the LR parse table. The LR parse table may be provided with a processing reference function for performing temporal relationship decision processing and semantic expression generation processing simultaneously with referring to the LR parse table. In this context, this analysis method is called a keyword lattice LR (KLLR) analysis method.
>
> ('768 patent, col. 5 line 62 – col. 6 line 12)

For this additional reason, Applicant believes that the semantic parser of claim 1 is not anticipated by the '768 patent.

In summary, for all of the foregoing reasons, upon closer inspection, the '768 patent does no anticipate each and every element of claim 1. Accordingly, Applicant respectfully requests removal of this rejection.

Claims 2-3

With respect to claims 2-3, the office action asserts that "Tsuboi teaches set of words and where 'The syntactic rewrite rules are written according to context free grammar or regular grammar'". These claims depend from claim 1, discussed in detail above and as such are believed to be patantable over the '768 patent. For example, the speech recognition section 2 of the '768 patent does not produce word paths. Applicant respectfully requests removal of these rejections.

Claims 4-6

With respect to claims 4-6, the office action asserts that "Tsuboi teaches where The recognition result of the speech recognition unit 2 is sent to a syntactic/semantic analysis unit 3 in the form of a series of temporally discrete words (i.e., word lattice)." However, this text of the '768 patent was discussed above in detail in relation to the phonetic search module. For the reasons stated above, these claims are not anticipated by the '768 patent. Applicant respectfully requests removal of these rejections.

Claims 7-18

As stated in the Office Action, claims 7-18 are analogous to claims 1-8 and are rejected for the foregoing reason. Accordingly, for the reasons set forth with respect to claims 1-7, Applicant requests removal of these rejections/

Claims 19

As to claim 19, the office action asserts that "Tsuboi teaches a syntactic/semantic tree (Fig. 1b, 20)." However, FIG. 1B of the '768 patent is not a an RSG in the form of a semantic tree – it's a sub-sentence. And FIG. 20 does not show a tree. Thus, Applicant believes this claim is not taught by the '768 patent and respectfully requests removal of this rejection.

Claims 20-22

As to claims 20-22, the office action suggests that "Tsuboi teaches a semantic expression evaluation (Fig. 17a)." Yet, as discussed above, the '768 does not teach a semantic evaluator interacting with an application program to determine a lunguistic result from the semantic data. Thus, Applicant believes these claims are not taught by the '768 patent and respectfully requests removal of these rejection.

Claims 23-30

With respect to claims 23-30, the office action suggests that "Tsuboi teaches a "word" searcher (3) as addressed above, coupled to a "grammar/dictionary" storage (4) also described above and configure to receive the phonetic stream of data and to generate series of words as function of syntactic information derived from the "grammar/dictionary" database (Fig. 17a)." Each of these elements was discussed above in detail with respect to claim 1. And for the reasons stated above, the Applicant believes that these claims are not anticipated by the '768 patent.

The Commissioner is hereby authorized to charge any additional fees under 37 C.F.R. §1.16 and §1.17 that may be required, or credit any overpayment, to our Deposit Account No. 50-1133.

Respectfully submitted,

David M. Mello, Reg. No. 43,799
McDermott, Will & Emery
28 State Street
Boston, MA 02109
Tel (617) 535-4037
Fax (617) 535-3800
E: dmello@mwe.com

Date: August, 30, 2004